BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

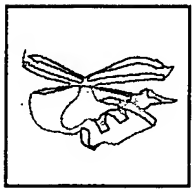
IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Automation of X-ray crystallography

Enrique Abola[1,2], Peter Kuhn[3], Thomas Earnest[4] and Raymond C. Stevens[1,2,5]

Structure-based biological discovery is entering a new era with the development of industrialized macromolecular structure determination pipelines. Intense, highly focused X-rays from integrated synchrotron radiation beam lines combined with significant advances in protein expression, purification, and micro-crystallization automation allow for the full streamlining of the traditionally tedious and time consuming process of determining the three dimensional structures of macromolecules.



Recent advances in macromolecular X-ray crystallography have laid a solid foundation from which a production pipeline optimized towards the determination of large numbers of structures can be constructed. Industrialization of the process will require the assembly of a linear series of high-throughput steps from target selection to structure interpretation, in which each step must be successfully completed before the next step is attempted (see Box 1 for a basic description of the steps involved in X-ray crystallography). It will also require the automation, miniaturization, parallelization, and optimization of existing crystallographic approaches. Improvements in the processes will have a significant impact on both the efficiency and the effectiveness of experimental techniques in structural biology. These developments are critical for structural genomics efforts as well as for rational drug discovery using high-throughput macromolecular crystallography. When implemented, these techniques should usher in a new era of structure-based biological discovery.

This review focuses on the process steps starting from the crystallization of protein samples to the collection and preliminary analysis of X-ray data. We describe each important task within these steps, identifying possible bottlenecks and efforts to overcome them primarily through automation. Other reviews in this issue focus on steps that occur before (protein production; see the article by Edwards and colleagues) and after (diffraction data analysis; see the article by Lamzin and Perrakis) those discussed here.

High-throughput protein crystallization

The crystallization step normally commences after the successful preparation of highly purified and soluble samples[1]. Attempts are then made to explore the phase space of a protein system by slowly moving the system towards a meta-stable supersaturated state. The actual crystallization conditions can at present not be determined *ab initio*, and the standard protocol is to optimize several solution variables (in a combinatorial approach) known to induce nucleation, such as pH, ionic strength, temperature, and specific concentrations of organic additives, salts and detergent. A complete exploration of this multi-dimensional condition space is impractical and reduced sampling approaches have been developed[2,6]. These methods have been successful, but are limited by the amount of protein available and the thoroughness with which this poorly defined multi-dimensional parameter space can be covered.

Automation of these crystallization processes was a focus at the 8th International Conference on the Crystallization of Biological Macromolecules (ICCBM8)[3]. New generation robots have been developed for the automated production of high-quality protein crystals at the Hauptman-Woodward Medical Research Institute, at the Protein Structure Factory in Germany, and by a collaboration between the Lawrence Berkeley National Laboratory (LBNL) Bioinstrumention Group, the Genomics Institute of the Novartis Research Foundation (GNF), and Syrrx, Inc. (a company devoted to high throughput structure biology)[4,5]. These integrated systems are capable of performing ~10,000 to >100,000 trials a day, using an expanded search of this multi-dimensional condition space[6]. Additionally, both success and failure data can accurately be accumulated for data mining and analysis for future improvements. Both the Hauptman-Woodward and LBNL/GNF/Syrrx system reduce the protein requirements by miniaturization of the experiments to nanoliter volumes. Imaging and analysis of these samples remains a challenge although several advancements have recently demonstrated they can be accomplished in a high throughput manner.

Crystal harvesting and storage for shipping

Once crystals are obtained, they have to be harvested and prepared for data collection. In most cases, structural genomics efforts will rely on synchrotron radiation facilities for efficient and rapid collection of diffraction data (see Box 2 for a basic description of synchrotron radiation). Therefore, samples must be packaged and stored for shipping. One breakthrough in macromolecular crystallography synchrotron data collection has been the routine freezing of protein crystals[7,8]. This is particularly important for multi-wavelength anomalous dispersion (MAD) data collection, where radiation decay can have a serious impact on the experiment. Flash-cooled crystals are stored at cryogenic temperatures and transported to data collection facilities, an approach described by Bob Sweet at Brookhaven National Laboratory as "Protein Crystallography by FedEx™"

[1]Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, USA. [2]Syrrx, Inc., San Diego, California 92121, USA. [3]Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, California 94305, USA. [4]Berkeley Center for Structural Biology, Physical Biosciences Division, MS6-2100, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [5]The Genomics Institute for the Novartis Research Foundation, San Diego, California 92121, USA. Correspondence should be addressed to R.C.S. *email: stevens@scripps.edu*

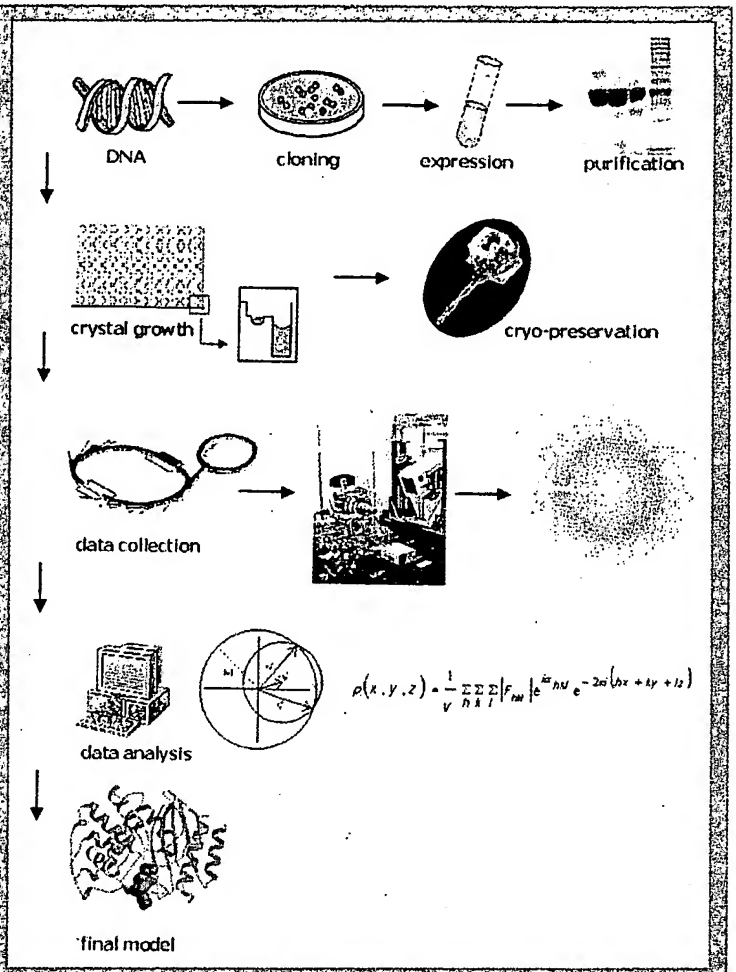
Box 1. Determining protein structures by X-ray crystallography

Protein structure determination involves the following steps: (i) preparation of highly purified protein samples (ii) crystallization (iii) measurement of diffraction data (iv) solving the phase determination problem using natural or artificially incorporated heavier atoms (v) phase and electron density calculations (vi) map interpretation and model building, and (vii) model refinement.

Incoming X-rays are diffracted by the electrons of the protein atoms to produce diffraction patterns from which a three-dimensional structure of a protein can be derived. In diffraction experiments, one is able to measure the intensity of the scattered beam but not the phases relating these scattered waves. X-ray crystallographic methods developed during the last few decades have focused on improving the measurements and recovering this phase information.

The diffraction experiment is simply the process of orienting a protein crystal in an intense X-ray beam and providing a mechanism for detecting most of the scattered beams. This is done with an integrating or photon counting device. Most modern detectors are multi-element charge coupled device (CCD) systems, which are capable of accurately measuring even weakly scattered X-ray beams. The majority of current diffraction data collection is done using the very intense and tuneable X-ray beams from synchrotron storage rings (see Box 2).

Two advances in X-ray crystallographic methodology developed over the past two decades, are cryo-freezing of protein crystals, which significantly reduces the radiation damage from the intense X-ray beams, and the development of the multi-wavelength anomalous dispersion (MAD) method. The MAD method relies upon the tuneability of synchrotron radiation to different wavelengths, to select for the absorption edge of metal atoms incorporated into protein crystals. This process facilitates the solving of the phase problem. In particular, incorporation of seleno-methionine to substitute for natural methionine residues has tremendously improved the ability to rapidly and routinely determine protein structures.



(ref. 9). Currently, these processes are done manually. Researchers work under a microscope to select a sample. The crystal is mounted into a cryo-loop, flash cooled at liquid nitrogen temperature, and subsequently stored in a shipping dewar. Routine handling of larger numbers of crystals will clearly require automation and resolution of other material handling issues, such as the ability to automatically identify each unique sample as it is mounted in the data collection station.

The task of harvesting crystals from multi-well plates represents a challenging automation problem. A prototype system was designed by Oceaneering Space Systems (Houston, Texas), under contract to the University of Alabama at Birmingham, for use in microgravity environments[10]. The system removes crystals and a small amount of mother liquor from the crystallization experiment. It isolates candidate crystals in a liquid bridge formed between two pipette tips, moves a robotic tip with a cryo-loop through the liquid bridge, and captures and freezes the crystal sample. As the system is designed to work on the space shuttle and international space station, a small robot with six degrees of freedom handles all manipulations. The robot is remotely controlled, allowing earth-based scientists to make decisions regarding crystal harvesting. This represents the first prototype from which subsequent systems can be developed using similar or different approaches.

A workshop, organized by scientists from the Advanced Light Source (ALS), the Advanced Photon Source (APS) and the Stanford Synchrotron Radiation Laboratory (SSRL), was recently held at Stanford University[11] to discuss different approaches to macromolecular crystal sample storage, mounting and characterization. Cryogenic storage, transport and retrieval systems require robotic sample handling at liquid nitrogen temperatures. At the meeting it was proposed that a standardized design will eliminate the need to change storage devices when using beam lines at different synchrotrons. Several projects are currently underway to develop and test prototypes. One example of a conceptual design from SSRL is depicted in Fig. 1, which shows a 96-pin cassette that holds 96 crystals at liquid nitrogen temperatures and the robotic system that extracts individual samples from the

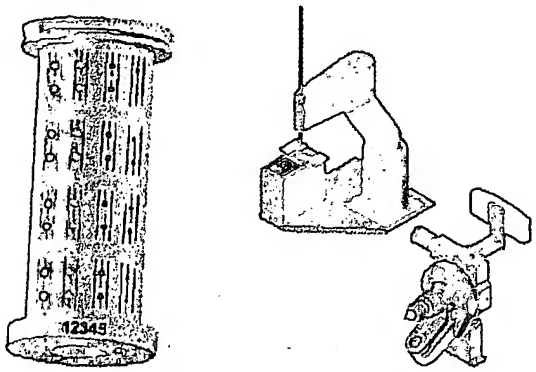


Fig. 1 On the left is the 96-sample cassette, which holds and retains flash-cooled crystals at liquid nitrogen temperatures. On the right is the conceptual design of a transfer robot that removes the flash-cooled sample from the cassette and transfers it at liquid nitrogen temperatures to mount it on the diffractometer. This system is the conceptual design of the SSRL SMB resource funded by the NIH-NCRR and DOE-BER.

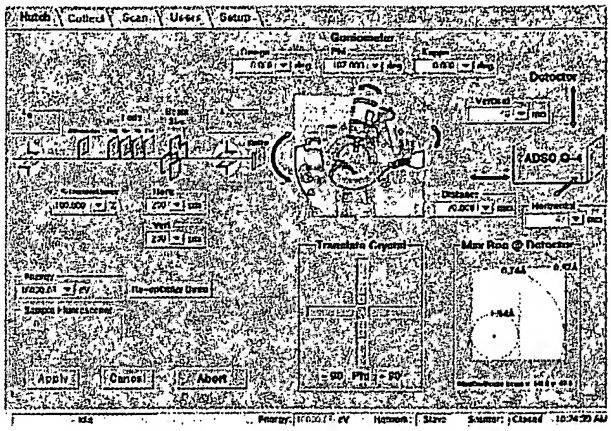


Fig. 2 'Hutch' window of BLU-ICE showing the intuitive interface that allows the experimenter to interact with the synchrotron beam line instrumentation. Complex hardware systems are integrated in the software package to allow ease of access for optimized design of experimental procedures. The photon energy can be changed by modification of a single number, but this change actually involves movement of several optical elements and table positions to new positions.

storage device and mounts them on a diffractometer at the synchrotron beam line. The cassette is bar-coded and each crystal location will be identifiable by a coordinate system, allowing for the unique identification of each sample.

Data collection at synchrotron beam lines

Synchrotron beam lines were initially developed as experimental stations, not for high throughput data collection. Therefore, data collection experiments currently require the user to enter the beam line hutch, mount and center a crystal, and exit/lock the hutch for every individual sample. These tasks result in a tremendous loss of valuable and scarce synchrotron beam time. Table 1 outlines estimates made by the Southeast Regional Collaborative Access Team (SER-CAT) group at APS of the amount of time used for a manual data collection procedure *versus* a potential automated procedure. In reality, no experimenter can keep up the optimal manual pace outlined in the table over a 24 hour period. Therefore, the throughput gain due to automation can be even more significant.

Automated approaches to data collection have become possible in recent years with the availability of reliable synchrotron sources, stable X-ray optics, rapid readout charge coupled device (CCD) detectors and high-precision diffractometers. Many of the synchrotron beam line development groups are currently implementing various automation systems that allow for mail-in crystallography and the development of research collaboratories to solve challenging problems in structural biology[11]. Advanced instrumentation and sophisticated software environments will enable full automation of synchrotron beam lines for high-throughput data collection to meet the goals of the integrated structure determination pipeline. These systems will allow the overwhelming number of crystallographic projects to be successfully pursued by many research groups. The data acquisition systems currently implemented by the National Synchrotron Light Source (NSLS) and SSRL are capable of remote operation and joint execution of experiments[12]. Fig. 2 shows the setup window of the system called Beam Line Unification in an Integrated Control Environment (BLU-ICE)[13], which is an example of a completely integrated software environment that functions as a distributed control system for crystallo-

Table 1 Task analysis showing typical time needed for data collection and estimated time savings with automation[1]

	Current manual estimate (seconds)	Automated estimate (seconds)
Open the hutch door and walk to the goniostat	25	0
Move the detector back	10	10
Retrieve the old sample	20	5
Mount a new sample	20	5
Align the new sample	90	45
Move of the detector back to data collection position	10	10
Set-up the interlock, leave the hutch and close the door	30	0
Wait for the end of the warning announcement	30	0
Open the photon shutter and return to the console	5	0
Beam stop alignment check	120	0
Start exposure	10	10
Total mounting time for each crystal	370	85
Repeat this for 3 crystals (needed to find a good crystal)	1,110	255
Data collection time	600	600
Total time for a data set	1,710	855

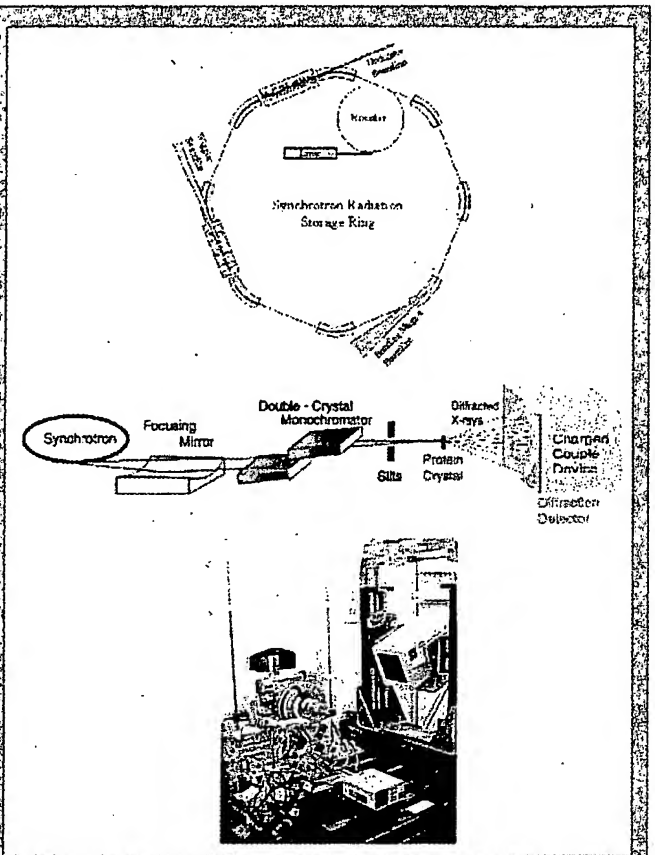
[1]Data collection times vary grossly and are crystal-specific. In reality, no experienced experimenter can keep up the optimal manual pace outlined in the table over 24 hours, nor are all users experienced at the same level. Table was kindly provided by B.C. Wang, principal investigator of the SER-CAT consortium at APS.

Box 2 Using synchrotron radiation for structural genomics

Synchrotron radiation is now indispensable for structural studies and will be the primary X-ray source for structural genomics efforts. Biologists started using synchrotron facilities in the early 1970s and by the 1980s, synchrotron radiation was helping to push the envelope by allowing crystallographers to work with increasingly larger molecules, such as virus particles and the intact ribosome.

Synchrotron radiation is produced when a high-energy particle beam changes direction in magnetic devices such as bending magnets, undulators or wigglers. The particle beams are either electron or positron beams that are generated by a linear accelerator (LINAC), further accelerated in a booster ring, and transferred into a storage ring to reach typical final energies of 1 GeV to 8 GeV. In the storage ring, the charged particles are bent around a polygonal path using magnetic fields produced by electrical coils or permanent magnets. These magnetic fields result in transverse acceleration of the particles, producing very intense, and highly parallel X-rays tangential to the particle beam that can be extracted and finely collimated to produce a focused beam that is extremely intense (orders of magnitude more intense than in-house X-ray sources) and tuneable. These two aspects make synchrotron-generated radiation an important tool for studying macromolecular structures.

The broad-spectrum radiation produced by synchrotrons is tuneable — it is possible to conduct studies anywhere in a typical wavelength range of 0.3–2.5 Å. This is unlike standard laboratory sources that produce X-rays at a particular wavelength. The X-rays can be selected from this broad spectrum using sophisticated X-ray optics, which makes it possible to easily conduct multiwavelength anomalous dispersion (MAD) studies on protein derivatives. Selecting lower wavelengths (0.3–1.0 Å) also minimizes the effects of absorption while retaining good intensity of diffraction.



A schematic of a generic synchrotron source and a beamline (top) and a picture of a standard experimental hutch (bottom) are shown. A typical synchrotron radiation source consists of a linear accelerator (LINAC), a booster ring, a storage ring and the beamlines. The sources of radiation are the magnetic devices that guide the electron path. These devices are typically bending magnets, wigglers, and undulators. Bending magnets are the 'arcs' of the polygon defining the synchrotron storage ring. Wigglers and undulators are installed in the straight sections; these devices consist of a series of alternating magnets causing the electron beam to 'wiggle' resulting in the emission of very intense X-ray beams with different spectral characteristics. The X-rays are guided and conditioned by using optical devices such as focusing mirrors and monochromators to project a focused, well-collimated and monochromatic beam on the sample, which is mounted on a diffractometer (middle). Finally, the experimental hutch contains the final beam conditioning system, sample mounting devices, and diffraction detector system.

graphic data collection. With BLU-ICE, users anywhere in the world can observe or have full control of all experimental instrumentation at a synchrotron hutch through a single, intuitive graphical interface. This control environment has eliminated the requirement for the experimenter to operate a beam line using multiple computer interfaces and/or mechanical devices. This unified access to all beam line instrumentation provides a clear path for further automation of other process steps.

Automated crystal mounting, viewing, and centering

Projects are currently underway to implement beam line automation leading to the time savings outlined in Table 1. Technical and process issues are being addressed to automate the mounting, centering and characterizations of crystals. During the Stanford workshop[11], Steve Muchmore from Abbott laboratories described an operational system that allows their group to collect multiple data sets, uninterrupted, on a standard laboratory rotating anode based system. What made this possible was their use of an automated sample changer and alignment system. The system consists of a three-axis manipulation robot fitted with a custom designed tool for crystal mounting, and a video camera used to obtain images of mounted crystals to perform the automated alignment procedures. These components have been integrated with a commercially available CCD detector system. The system includes a sample storage rack containing up to 63 pre-frozen crystals mounted on cryoloops. The racks are stored in an open dewar in which a layer of liquid nitrogen gas is used to reduce icing. Although constructed for use with a non-synchrotron X-ray source, the design and implementation allow for integration into synchrotron beam lines.

A second system is currently being developed specifically for synchrotron beam line use. The bioinstrumentation group at LBNL, in collaboration with ALS and researchers at GNF as well as Syrrx are developing a first generation automated crystal

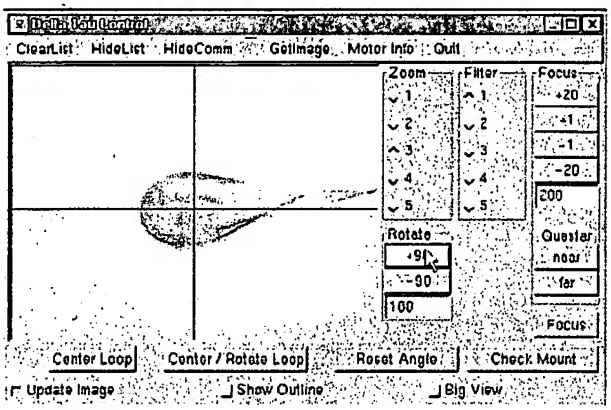


Fig. 3 Autocentering of protein crystals at the Advanced Light Source. By clicking on the crystal location, the sample is automatically centered in the X-ray beam.

recognition and alignment system designed to work at synchrotron beam lines. The system employs a task-specific design in which a standard cryovial is moved to a goniometer. An electromagnet holds the magnetic cap and sample while the vial is replaced into a liquid nitrogen dewar. Automation of crystal alignment using an intelligent software system is currently under development (Fig. 3). When operational, this system will allow for the screening and data collection of several hundred crystals a day.

Finally, the handling of very small crystals is being made possible by the development of newly designed microcrystal diffractometers such as the Grenoble EMBL/ESRF microfocus beam line ID13 (refs. 14,15). The ID13 beamline is designed to allow maximum precision positioning combined with ease of usage necessary for handling very small crystals (down to 5 μm). The design centers around a parallax-free video microscope used to view both the sample and a beam localization scintillator. A motorized sample centering system, ϕ-axis positioning tables, beam definition aperture for beam sizes of 10–200 μm, and beamstop allow for computer-aided automation. Currently, sample loading is done manually and crystal alignment is performed *via* a graphical user interface. Under development is an automated sample changer for pre-frozen crystals capable of handling up to 24 samples stored in cylindrical racks. Each of these samples can be transferred automatically to/from the goniometer axis. Use of the micro-diffractometer allows for the use of very small crystals and also allows for diffraction from localized areas of twinned crystals. These features are desirable for high-throughput studies, especially for those using reduced protein volumes in crystallization experiments that are expected to produce smaller crystals.

Diffraction quality analysis, indexing, and data collection

Most structure determinations require the screening of multiple crystals in order to select the most suitable sample with regard to diffraction quality and orientation. The number of crystals used depends upon the complexity of the project; for example, very few test crystals are required for determination of additional structures, such as of a mutant protein or a protein in complex with an inhibitor, but hundreds are required for more difficult membrane proteins or multi-domain complexes. The mounting/unmounting and initial crystal characterization represent time consuming steps, which can be made more efficient by automation. Robotic handling and data maintenance reduce the chance of human error, which can be introduced in manual methods by the required repeated handling of each sample, as well as the data assessment of many different samples. Crystallographic software development groups are currently developing and testing software to evaluate diffraction quality and establish optimal data collection strategies using currently available beam line capabilities (such as multi-circle diffractometers, high-speed CCD detectors, and detector off-set).

Conclusion

Industrialization of the structure determination process will increase final throughput numbers and lead to an enhanced success rate and higher quality of derived three-dimensional models. It will result in the establishment and population of databases that contain accurate information on all process steps, including both successes and failures. These databases will contain information related to target selection, protein expression/crystallization, data collection and refined structures. Systematic analysis of these databases will display trends and provide answers to questions for rational experimental design so that there will ultimately be less reliance on anecdotal stories that historically have directed experiments on an almost trial and error basis.

Acknowledgments

We appreciate the use of material presented at the SSRL workshop on automation development for synchrotron beam lines held May 11-12, 2000, and information from the 8th International Conference on the Crystallization of Biological Macromolecules (ICCBM8) held in Sandestin, Florida; May 14-19, 2000. In particular, we appreciate the material from S. Muchmore and B.C Wang that was presented in this review. We are grateful to M. Scott for careful reading and helpful discussions of this review. The SSRL and ALS structural biology programs are supported by NIH (NCRR, NIGMS) and the DOE (BES, BER).

Associations with structural genomics

E.A. is a consultant for Syrrx. P.K. is a member of the Joint Center for Structural Genomics (JCSG) between TSRI, UCSD, and Stanford. T.E. is Head of the Macromolecular Crystallography Facility at the Advanced Light Source. R.S. is a member of the JCSG and a consultant for the Genomics Institute for the Novartis Research Foundation (GNF). R.S. is also a founding scientist of Syrrx, a company focused on high throughput protein crystallography for drug discovery.

1. Stevens, R.C. *Structure* **8**, R177–R185 (2000).
2. Janacarik, J., Kim, S.H.J. *J. Appl. Cryst.* **24**, 409–411 (1991).
3. 8th International Conference on the Crystallization of Biological Macromolecules (ICCBM8), Sandestin, Florida; May 14–19 (2000).
4. Luft, J. et al. 8th International Conference on the Crystallization of Biological Macromolecules (ICCBM8), Sandestin, Florida; May 14–19 (2000).
5. Saenger, W. & Lehrach, H. 8th International Conference on the Crystallization of Biological Macromolecules (ICCBM8), Sandestin, Florida; May 14–19 (2000).
6. Carter, C., Jr. In *Crystallization of nucleic acids and proteins. A practical approach 2nd Edition.* (eds Ducruix, A. & Giege, R.) 75–118 (Oxford University Press; Oxford, UK; 1999).
7. Rodgers, D.W. *Methods Enzymol.* **276**, 183–203 (1997).
8. Garman, E.F. & Schneider, T.R. *J. Appl. Crystallogr.* **30**, 211–237 (1997).
9. Sweet, R.M. *Nature Struct. Biol.* **5**, 654–656 (1998).
10. http://www.oceaneering.com/oii/view/rob0005.htm. Crystal Preparation Prime Item (CPPI).
11. http://smb.slac.stanford.edu/robotics -Workshop on techniques for automated mounting, viewing and centering pre-cooled protein crystals; Stanford Synchrotron Radiation Laboratory, Palo Alto, California: May 11–12 (2000).
12. Skinner, J. M. & Sweet, R. M. *Acta Crystallogr. D* **54**, 718–725 (1998).
13. http://smb.slac.stanford.edu/blu-ice; BLU-ICE and the Distributed Control System at SSRL.
14. Cusak, S. *et al. Nature Struct. Biol.*, **5**, 634–637 (1998).
15. Perrakis, A. *et al. Acta Crystallogr. D* **55**, 1765–1770 (1999).